The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/2632-7627.htm

Education and economic growth: an empirical analysis of nonlinearities

Laura Marquez-Ramos
Department of Economics, Universitat Jaume I, Castellón, Spain and The University of Adelaide Institute for International Trade, Adelaide, Australia, and
Estefanía Mourelle
Department of Economics, Universidade da Coruna, A Coruna, Spain

Received 17 October 2017
Revised 31 May 2018
Accepted 5 October 2018

Abstract

Purpose – Might a country's economic growth performance differ depending on the evolution of its human capital? This paper aims to consider education as a channel for human capital improvement and then for economic growth. The authors hypothesize the existence of a threshold for education, after which point the characteristics of economic growth change.

Design/methodology/approach – To address this question, the authors turn from a linear framework to a nonlinear one by applying smooth transition specifications.

Findings – This empirical analysis for Spain points to the existence of nonlinearities in the relationship between education and economic growth at country level, for both secondary and tertiary education. Next, as different patterns emerge in different regions, the authors provide a regional analysis for a number of representative Spanish regions. The results show that both secondary and tertiary education matter for economic growth and that nonlinearities in this relationship should be taken into account.

Practical implications – What is learnt from using Smooth Transition Regression models for the education-economic growth link is that the educational level of the population can be understood as a source of nonlinearities in the economic activity of a country (and of a region). Thus, depending on national and regional educational levels, economic growth behaves differently.

Originality/value – Although the importance of nonlinearities has been identified, linearity is usually assumed in this field of the literature. This paper calls into question the linearity assumption by using time series techniques for 1971-2013 in Spain, an OECD country, and testing whether the results at country level hold for different regions within Spain as a robustness check.

Keywords Education, Economic growth, Nonlinearities

Paper type Research paper

1. Introduction

This paper analyses the relationship between education and economic growth in Spain. It makes a contribution to the literature by analysing the existence of nonlinearities in

JEL classification – C32, C59, I25, R11

this relationship. The direction of the causality that we account for is from education to economic growth. In some cases, an increase in education might have a positive (or large) effect on GDP growth and, in other cases, the effect might be negative (or small). Because the effect of education might differ in different stages, nonlinearities gain importance in the analysis of the relationship between education and economic growth. Specifically, we hypothesise that a country's economic growth performance differs depending on the evolution of its human capital and hence, on its educational achievement.

Education is a key determinant of economic well-being and increases the human capital inherent in the labour force of a country (Hanushek and Wößmann, 2010). In this paper, we start from the assumption that a country with a population that is capable of exploiting new knowledge will perform better; as such, education is considered a channel for economic growth since it constitutes an intrinsic mechanism for knowledge absorption. We focus on both secondary and tertiary education, and use time series techniques to analyse the existence of nonlinearities.

Cohen and Levinthal (1990) introduced the concept of absorptive capacity, which is the ability to recognise the value of new, external information, to assimilate it and to apply it. Education improves this ability and so absorptive capacity can explain why a nonlinear relationship between education and growth might be expected, i.e. why a country's economic growth performance might differ depending on its educational achievement[1].

Our research has a bearing on the literature on education and economic growth. Previous studies have analysed the causal effect of education on countries' economic performance using a variety of measures and quantitative tools for cross-country, time series and panel data (Hanushek and Wößmann, 2010). From the literatures on growth, education and development, we know that the difference between having a primary education and being illiterate has a relatively higher impact on individual wages than having a secondary or tertiary education (i.e. there are decreasing returns to education over certain thresholds). Thus, nonlinearities exist at the micro level, and the macro literature has also explored diverse effects. For example, previous studies have revealed country heterogeneity, showing that estimated returns to education are, in general, higher in developing countries than in developed countries (Duflo, 2001).

Moreover, there is a very large macro (cross-country and panel) literature analysing the relevance of education and economic growth, which has focused on different types of education. Within this research line, Sunde and Vischer (2015) find that there is a weak empirical effect of human capital on economic growth in existing cross-country studies, which is partly due to inappropriate specifications. Previously, Kalaitzidakis *et al.* (2001) found a nonlinear effect of human capital on economic growth. Kalaitzidakis *et al.* (2001) point out that their evidence "is consistent with the theoretical suggestion that there exist threshold levels of human capital and the growth experience of a country may well differ according to which side of the threshold it finds itself in" (p. 251).

The importance of nonlinearities is illustrated by Krueger and Lindahl (2001). These authors divide their sample of countries into three subsamples based on their initial human capital endowment, finding a positive association only in the subsample of countries with the lowest educational level. For the group of countries in the middle of the education distribution, they either find no effect of education on growth or a negative relationship. The relationship is also negative in countries with a higher educational level. In other words, the authors find that "the relationship is inverted-U shaped, with a peak at 7.5 years of education" (p. 1130). Given that the average years of education in OECD countries was 8.4 in the year of the study, these authors point out that "the average OECD country is on the downward-sloping segment of the education-growth profile" (p. 1130).

Although the importance of nonlinearities has been identified, linearity is usually assumed in this field of the literature (Self and Grabowski, 2004; Cohen and Soto, 2007; Afzal *et al.*, 2011; Armellini, 2012; Jalil and Idrees, 2013). This paper calls into question the linearity assumption by using time series techniques for 1971-2013 in Spain, an OECD country, and testing whether the results at country level hold for different regions within Spain as a robustness check[2].

When analysing the relationship between education and economic growth, it must be acknowledged that Spain is one country where *over*-education might be a significant issue. Given the high unemployment rate in Spain (resulting in underemployment), the impact of education on economic growth might be expected to be close to zero. That is, increasing educational levels might not necessarily be associated with higher economic growth in Spain. Regarding to the specific hypothesis on the relationship between education and economic growth in Spain, it is worth noting that the answer to this research question is ambiguous, i.e. this relationship can be positive, negative or non-significant.

In our empirical analysis, we incorporate a number of explanatory variables that are closely related to both education and economic growth, as they also affect this relationship; based on the related literature, we account for the potential effects that physical capital, labour force and public expenditure on education may have on the economic activity. We are able to provide evidence of a positive correlation between education and economic growth, as well of the existence of a nonlinear relationship at both country and regional level. This analysis helps us to shed light on the question posed by the existing literature about the positive, although not yet universally accepted, impact of education on economic growth.

The rest of the paper is organised as follows. Section 1 presents the methodology employed. Section 2 details the empirical results obtained at country level, while Section 3 presents the results at regional level. Finally, Section 4 concludes.

2. Methodology

One common assumption when analysing the relationship between economic growth and education is that of linearity. However, this assumption would mean that the parameters in this relationship do not change over time[3]. Thus, we go beyond the conventional use of linear models by exploring nonlinear specifications that might more accurately explain the relationship between education and economic growth.

Threshold regressions are one of the most common specifications used for reflecting regime changes. As these models consider abrupt changes between regimes, we adopt smooth transition (ST) specifications and, in particular, the more generalised smooth transition regression (STR). In this case, the variable is assumed to vary between two extreme regimes and the smoothness of the transition is estimated from the data. This framework provides us with three main advantages. First, it offers more flexibility in studying the dynamics of the relationship between economic growth and education, not only enabling a wide range of nonlinear behaviours to be described, but also allowing for a continuum of intermediate regimes. Second, it allows us to analyse the potential existence of a threshold that determines the behaviour of economic growth according to education. Finally, it permits the incorporation of exogenous variables in addition to the endogenous structure[4].

Let y_t be a stationary ergodic process and, without loss of generality, only one exogenous variable x_t. The model is given by:

$$y_t = w_t^{'}\pi + \left(w_t^{'}\theta\right)F(s_t;\ \ \gamma, c) + \ u_t, \tag{1}$$

where $w_t = (1, y_{t-1}, \ldots, y_{t-p1};\ x_t, x_{t-1}, \ldots, x_{t-p2})'$ is a vector of regressors; $\pi = (\pi_0, \pi_1, \ldots, \pi_p)'$ and $\theta = (\theta_0, \theta_1, \ldots,\ \ \theta_p)'$ are parameter vectors ($p = p_1 + p_2 + 1$); and u_t is an error process, $u_t \sim$ Niid (0, σ^2). Likewise, F(.) is a transition function customarily bounded between 0 and 1, implying that the STR coefficients vary between π_j and $\pi_j + \theta_j$ ($j = 0, \ldots, p$), respectively. The regime at each t is determined by the transition variable, s_t, and the associated value $F(s_t)$; the transition variable can be a lagged endogenous variable, an exogenous variable or just another variable.

In this study, we select the logistic transition, as it appears to be the most suitable for describing the relationship between education and economic growth, as there is no reason to assume that a positive and a negative variation in a country's educational level will have a similar effect on its economic activity. The slope parameter γ defines the smoothness of the transition from one regime to the other, such that the higher the value, the more rapid the change; and the location parameter c indicates the threshold between the two regimes. The logistic case involves that $F(-\infty) = 0$ and $F(\infty) = 1$; this means that the extreme regimes (related to $F = 0$ and $F = 1$) are associated with s_t values far above or below c, where dynamics may be different.

Regarding to the modelling procedure, we follow a well-established strand of the literature that focuses on an extensive search of STR models. The validation of the estimated models is the core part of the procedure. Accordingly, we apply the three specific tests designed by Eitrheim and Teräsvirta (1996) for STs as misspecification tests. In terms of diagnostic statistics, we employ the adjusted coefficient of determination and we pay particular attention to the variance ratio of the residuals from the nonlinear model and the best linear regression estimated by ordinary least squares (OLS), as it provides relevant information on the explanatory power of both specifications. Finally, these evaluation tests are completed with an analysis of the estimated residuals to describe the behaviour of the STR models more in depth. Appendix 1 provides a detailed explanation of the model, as well as of the modelling procedure.

It is worth mentioning that, beyond the better econometric properties shown by STR models compared to linear ones, the use of nonlinear models allows us for characterising the "economic growth-education" dynamics in a more complete manner than linear models. It is well-known that the linear model is very restrictive; were the relationship between economic growth and education linear, a change in education would be associated to a particular change in economic growth, independently of the level of education reached. Under the linear framework, only one state is possible. However, scenarios with multiple states seem much more realistic.

In our case, and following the related literature, it makes sense to think that the economy might behave differently depending on the educational level of the population. This asymmetric behaviour cannot be captured by a linear model, so the search for nonlinearities becomes a crucial question. In this sense, ST models allow us for defining different regimes regarding the educational level so that we can determine whether they affect the economic activity in a different manner. Thus, these nonlinear models give us more information on the dynamics of the relationship education-economic growth than linear models. For instance, we are able to know how rapidly economic growth reacts to changes in education.

3. Empirical analysis

3.1 Data and descriptive analysis

We use data for Spain and carry out both an aggregate study at country level and a disaggregated analysis for a group of regions to test the robustness of the main results. Table AI in Appendix 2 displays the variables used in the empirical analysis, their definitions and sources.

GDP is traditionally used as a measure of economic activity. We proxy education by using enrolment ratios for both secondary and tertiary education. We define one proxy for labour force for each educational level; specifically, we calculate the proportion of the active population with secondary (and with tertiary) education over the total active population in Spain. Physical capital might exert a relevant direct effect on economic growth[5]. Thus, as a proxy for physical capital, we use real Gross Fixed Capital Formation. Finally, we use the government expenditure on education (as a percentage of GDP)[6]. The sample spans the period from 1971 to 2013 and the data are annual.

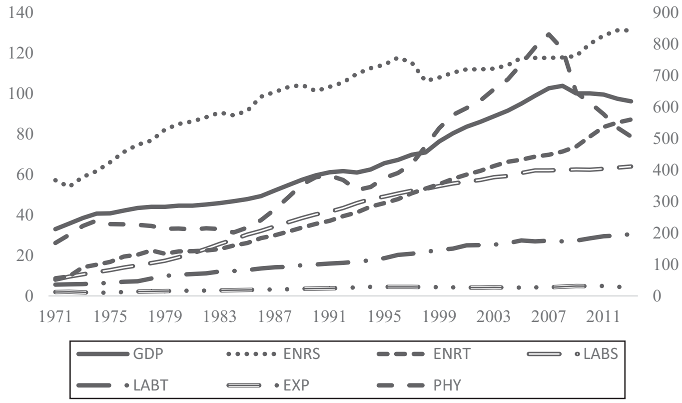
To provide an overview of the selected variables, we first provide brief descriptive statistics in Table I, then display their evolution in Figure 1.

Looking at Table I and Figure 1, we can highlight several features. The education variables show a positive trend over time. This is especially remarkable for tertiary education, which increases at a higher rate than secondary education; the latter even registers negative variations in a number of periods. Higher education shows a steady evolution, with important increases in the mid-80s, mid-90s and late 2000s. In general, all variables display an upward trend in recent decades, in particular after the end of the Franco dictatorship and the economic crisis of the 1970s. Thus, the mid-1970s is a crucial period for public government expenditure on education, with the proportion of the labour force with tertiary education starting to rise, and an almost continuous upward trend in the labour force with secondary education. In the 1980s, physical capital registers a substantial increase, which continues until the breakout of the last economic crisis. The upward trend in GDP is accompanied by a similar trend in the remaining variables. It is therefore challenging to isolate the effect of a causal relationship between economic growth and education using a time series approach.

	GDP	ENRS	ENRT	PHY	LABS	LABT	EXP
Mean	62	97.42	36.70	363.81	34.37	15.49	3.31
Median	61.03	105.37	39.33	368.19	43.40	16.26	3.85
Standard deviation	22.66	20.87	22.95	197.09	19.04	7.99	1.02
Kurtosis	−1.33	−0.39	−1.13	−0.90	−1.41	−1.35	−1.18
Skewness	0.37	−0.70	0.34	0.61	−0.32	0.05	−0.49
Minimum value	32.96	53.84	8.67	167.92	8.00	5.56	1.46
Maximum value	103.69	131.09	87.07	830.06	63.87	30.46	4.87

Notes: GDP: Gross domestic product; ENRS: enrolment ratio secondary education; ENRT: enrolment ratio tertiary education; PHY: physical capital; LABS: labour force with secondary education; LABT: labour force with tertiary education; EXP: government expenditure on education. All variables are in levels
Source: Own elaboration

Table I. Descriptive statistics for the selected variables (in levels)



Notes: Left axis: GDP, Gross Domestic Product; ENRS, enrolment ratio secondary education; ENRT, enrolment ratio tertiary education; LABS, labour force with secondary education; LABT, labour force with tertiary education; EXP, government expenditure on education. Right axis: PHY, physical capital. All variables are in levels
Source: Own elaboration

Figure 1. Evolution of the selected variables over time (in levels)

Before focusing on the relationship between economic growth and education, a statistical processing of the information is required, to carry out the suitable transformations of the variables. First, we use a logarithmic transformation of the original data; namely, a particular type of Box–Cox transformation customarily used to solve problems of nonstationarity in variance. Second, one assumption of ST models is that all the variables involved in the study must be stationary.

In this paper, we have applied the Ng and Perron (2001) unit root tests to analyse the order of integration of our variables. These authors propose several unit root tests (MZα, MZt, MSB and MPt) with the aim of improving the performance of existing ones, in particular regarding size and power features. In this paper, the Ng–Perron tests include intercept and linear trend as deterministic components, and the lag length has been selected by means of the modified Akaike Information Criterion proposed by the authors. The null hypothesis is the existence of unit roots. The test statistics are displayed in Table AIII (Appendix 3). Considering the asymptotic critical values defined in Ng and Perron (2001), all the variables analysed in the study are unit root processes. These results point to the need to apply regular differences in all (the logarithms of) our variables in the regression analysis.

3.2 Main results

Four explanatory variables are initially included in the regression analysis: education, physical capital, labour force, and public government expenditure on education. We consider up to two lags of the variables, so as to account for the effect of their most recent history on economic growth. In our paper, linearity tests involve working with a great number of cross products and our sample size is not very large, so the potential results should be taken with caution (Teräsvirta, 1998; and Skalin and

Teräsvirta, 2002). As the results would not be conclusive for the two educational levels, the modelling strategy will be based on the abovementioned extensive search of STR models and the core part of the modelling process will be at the validation stage[7].

The starting point of the modelling strategy consists of finding out the linear specification that would best describe the behaviour of the series under study. OLS estimation is carried out; all parameters are introduced initially, but those that are not significant at the 0.05 level are successively excluded. The next stage focuses on the estimation of the nonlinear models. We achieve valid STR specifications for secondary and tertiary education.

The traditional modelling cycle determines the transition variable by carrying out tests on all explanatory variables; however, following the literature, it can also be selected on the basis of the rationale for the analysis or the economic theory (Cheikh, 2012). As our aim is to study the effect of education on economic growth considering nonlinearities, the enrolment ratio for secondary and tertiary education is chosen as the transition variable. This decision is thus made on the basis of testing a theoretical hypothesis. In any case, any inadequacy arising in the first stages is likely to be later revealed in the modelling cycle (Teräsvirta, 1998). Thus, this variable plays a dual role in the explanation of the economic growth dynamics. On the one hand, it represents the source of nonlinearities in the evolution of economic growth; on the other hand, it is also a determining factor of the dynamics of GDP growth. The estimated linear and STR models are presented in Table II, along with several diagnostic statistics and misspecification tests.

The estimated models for secondary and tertiary education are reflected in columns one to four, taking into account the explanatory variables listed in column (0). For each educational level, the first column (Columns 1 and 3) presents the linear model and the second column is for the STR specification (Columns 2 and 4). With respect to the nonlinear models, the upper part of the table presents the regression coefficients (joint with their standard errors) corresponding to the "lower regime" ($F = 0$); the regression coefficients for the "upper regime" ($F = 1$) are the result of summing up the values in the upper and the lower parts of the table (i.e. interactions and no-interactions with F), as commented in Section 1. The smoothness of the transition between regimes is given by the slope parameter γ and c represents the estimated threshold.

Regarding the results of the linear specification, we might expect to find a negative or a positive association between higher education and economic growth in Spain. It is worth mentioning that reverse causality might be present. On the one hand, when GDP grows steadily at the same time as rising youth employment (as in the boom stage from 1996 to 2008), the opportunity cost of investing in education increases and the demand for post-compulsory studies, such as college, declines. On the other hand, when the crisis hits and the probability of a young person having a job falls abruptly, the opportunity cost of investing in tertiary education is reduced and the demand for higher education increases. This establishes not only an association between GDP growth (effect) and enrolment education (cause), but also an association between GDP growth (cause) and enrolment education (effect). The direction of the causality that we account for is from education to economic growth, but in some stages, an increase in education might have a positive (or large) effect on GDP growth and, in other stages, the effect might be negative (or small)[8]. Because the effect of education on growth might differ in different stages, nonlinearities gain importance in the analysis of the relationship between education and economic growth.

(0) Dependent variable: economic growth (GDP_t)	Secondary education Linear model (1)	STR model(2) Transition variable: $ENRS_{t-1}$	Tertiary education Linear model (3)	STR model (4) Transition variable: $ENRT_{t-2}$
"Lower" regime (F = 0)				
GDP_{t-1}	0.25 (0.10)	−1.31 (0.26)	0.28 (0.08)	0.66 (0.15)
GDP_{t-2}	0.19 (0.09)	1.43 (0.25)		0.32 (0.14)
$ENRS/T_t$	0.15 (0.06)		0.09 (0.03)	0.18 (0.06)
$ENRS/T_{t-1}$				−0.50 (0.12)
$ENRS/T_{t-2}$				
PHY_t	0.25 (0.03)		0.23 (0.03)	0.23 (0.03)
PHY_{t-1}		0.56 (0.07)		
PHY_{t-2}		−0.48 (0.13)		−0.18 (0.07)
$LABS/T_t$		−0.85 (0.32)		0.07 (0.03)
$LABS/T_{t-1}$		1.07 (0.36)	0.12 (0.04)	0.17 (0.05)
$LABS/T_{t-2}$		0.16 (0.09)		
EXP_t				
EXP_{t-1}	0.07 (0.03)			0.21 (0.08)
EXP_{t-2}				
Interactions with $F(s_t)$				
$F(s_t)\times GDP_{t-1}$		3.97 (1.38)		−1.35 (0.28)
$F(s_t)\times GDP_{t-2}$		−2.66 (0.80)		
$F(s_t)\times ENRS/T_t$				−0.34 (0.12)
$F(s_t)\times ENRS/T_{t-1}$				0.67 (0.14)
$F(s_t)\times ENRS/T_{t-2}$				
$F(s_t)\times PHY_t$		0.56 (0.14)		
$F(s_t)\times PHY_{t-1}$		−1.51 (0.48)		0.18 (0.07)
$F(s_t)\times PHY_{t-2}$		0.83 (0.27)		0.22 (0.07)
$F(s_t)\times LABS/T_t$		1.26 (0.53)		0.07 (0.03)
$F(s_t)\times LABS/T_{t-1}$		−1.76 (0.62)		
$F(s_t)\times LABS/T_{t-2}$				
$F(s_t)\times EXP_t$				0.08 (0.04)
$F(s_t)\times EXP_{t-1}$				−0.15 (0.08)
$F(s_t)\times EXP_{t-2}$		−0.08 (0.04)		
γ		2.17 (1.00)		66.65 (375.65)
c		0.015 (0.01)		0.052 (0.00)
$\overline{R}^2$	0.78	0.92	0.78	0.87
s^2/s^2_L		0.25		0.33
AUTO		2.89 (0.08)		3.52 (0.06)
NL		2.60 (0.23)		2.50 (0.32)
PC		1.40 (0.38)		1.09 (0.55)

Notes: Country-level analysis. GDP: Gross Domestic Product; ENRS: enrolment ratio secondary education; ENRT: enrolment ratio tertiary education; PHY: physical capital; LABS: labour force with secondary education; LABT: labour force with tertiary education; EXP: government expenditure on education. All variables are first differences of the logarithmic transformation. Values after regression coefficients are SEs of the estimates; $\overline{R}^2$ is the adjusted coefficient of determination; s^2/s^2_L is the variance ratio of the residuals from the STR model and the linear regression; AUTO is the test for residual autocorrelation of order 2; NL is the test for no remaining nonlinearity; PC is the general parameter constancy test. Numbers in parentheses after misspecification tests statistics are *p*-values
Source: Own elaboration

Table II.
Estimated models

In the linear model, both secondary and tertiary education present a positive and significant association with economic growth. According to the obtained results, a 10 per cent increase in secondary education leads to a 1.5 per cent increase in economic growth, while an increase of 10 per cent in tertiary education increases economic growth by 0.9 per cent, *ceteris paribus*.

Regarding the model including secondary education, the nonlinear model shows that economic growth depends on its own recent history (GDP), as well as on physical capital (PHY), labour force (see LABS/T in column 2) and public government expenditure on education (EXP). However, secondary education does not seem to be a significant factor determining economic growth, according to the final explanatory variables included in the model (see ENRS/T in column 2). Turning to tertiary education, results indicate that the estimated STR parameters are significant: recent history of economic growth has a significant effect on its current state (GDP), as do tertiary education (see ENRS/T in column 4), physical capital, labour force (see LABS/T in column 4) and government expenditure on education (the latter is not relevant for explaining economic growth in the linear specification). In contrast to secondary education, tertiary education is a significant determinant of economic growth dynamics in Spain (see ENRS/T in column 4).

Nonlinear models point to asymmetric effects of the different variables on economic growth. As shown in Table II, most variables exhibit a different sign when the variation in the enrolment ratio exceeds the location parameter (c equals 1.5 per cent for secondary and 5.2 per cent for tertiary education) and when it is below this threshold. This is proven by comparing the coefficients for the lower regime ($F = 0$) and those for the upper regime ($F = 1$); the latter are the results of summing up the coefficients for the lower regime and those for the interactions with the transition function.

Thus, for relatively large levels of secondary education enrolment (upper regime), we expect an overall negative effect of investment and labour on economic growth, but overall positive when the enrolment ratios are relatively low. When focusing on tertiary education, we observe how labour has always a positive effect (regardless of the regime) on economic growth. Physical capital affects in a positive manner especially when tertiary education enrolment is relatively high (upper regime). Interestingly, if the country shows large tertiary education enrolment ratios, their increase strengthens the role of education, as the economic activity improves (see upper regime); in case tertiary education enrolment were relatively low, its increase does not help economic growth in an overall sense (note the opposite signs for ENRS/T in moments *t* and t-1 in the last column). However, it should be noted that the interpretation of ST models merely based on the estimated coefficients can be misleading, as they do not completely account for the dynamics (Mejía-Reyes *et al.*, 2010).

As obtained for the linear model, results of the estimated coefficients obtained with the nonlinear specification indicate that higher education is a significant determinant of economic growth in Spain; however, in this specification, secondary education is not statistically significant[9]. Interestingly, both educational levels play a role in generating asymmetric effects on economic growth. In other words, while tertiary education plays the double role of being a transition and an explanatory variable, secondary education acts as a transition variable only. In any case, this result does not diminishes the relevance of secondary education on economic growth, as the behaviour of economic activity is determined by the regime education is.

Figure A1 in Appendix 3 depicts the estimated transition functions[10]. Regarding the secondary education case, the transition between regimes is logistic and smooth, according to the value of γ, and the delay is one year. Regarding the value for c (the estimated threshold), the Spanish economic activity will exhibit different dynamics when the enrolment ratio rises above 1.5 per cent (upper extreme regime) than when it goes at slower pace, that is, below that threshold (lower extreme regime). This function shows wide variation, allowing more flexibility in the dynamics of economic growth. The estimated value for the threshold is very close to the enrolment ratio mean (2 per cent), so that the left side of the logistic function contains fewer observations (66 per cent of the total).

With respect to tertiary education, we can define a lower extreme regime, ranging from negative growth to 5.2 per cent variation in the enrolment ratio (the estimated threshold, c), and an upper extreme regime, for variations greater than 5.2 per cent. This value is remarkably close to the mean of the transition variable (5.7 per cent), so there is a near equal distribution between the left and right sides of the logistic function. In this case, regime changes are extremely rapid (due to the large value of γ, which equals 66.65), a fact that suggests a need for threshold specifications and strengthens the importance of using STR models[11].

According to the estimated values of the slope parameter (γ) in Table II, economic growth appears to evolve more rapidly from one extreme regime to the other when considering tertiary education than secondary education, that is, economic growth shows a more immediate reaction to shocks in tertiary education than to shocks in secondary education. As stated in Section 1, the larger the value for γ, the more rapid the transition between regimes (γ equals 2.17 in secondary education and 66.65 in tertiary education). A possible cause may be the importance traditionally attributed in Spain to the pursuit of a university degree.

Following the evaluation tests, there are no indications of misspecification in the STR models[12], so we conclude that the proposed STR is adequate. Finally, a comparison analysis of the estimated residuals is carried out, comparing the residuals of the linear and nonlinear specifications for secondary and tertiary education. Figure A2 in Appendix 3 depicts the differences, in absolute values, between the residuals from the linear model and the STR specification over time. In both cases, the residuals of the STR model are lower (in absolute values) than those of the linear regression; in particular, the nonlinear model reduces overall the highest residuals of the linear specification in a reasonable way. Moreover, differences between residuals are mainly positive, thus supporting the use of the STR specification. These signs are a further indication of the superior performance of the nonlinear model relative to the linear model.

4. Results at regional level

To test whether results are robust to the degree of territorial aggregation chosen, in a first step, we construct a classification matrix to choose a representative sample of Spanish regions for the analysis. We classify regions by regional income per capita and regional government expenditure per capita. The information in Table AII in Appendix 2 is used to determine whether the different Spanish regions are above or below average in terms of both income and government expenditure on education. For example, when income per capita in a specific region is above the average over the analysed period, then this region is considered to be in a different cell of the classification matrix than a region that is below the average. A representative region is selected from each of the four groups of regions constructed (see Table AII, Appendix 2). We select four representative Autonomous Communities: Castilla y León, Cataluña, Comunidad

Valenciana and País Vasco. In three of these regions (Castilla y León, Cataluña and País Vasco), the average public expenditure on fundamental public services per capita for 2009-2015 was above the average of the communities with a common system (i.e. all Autonomous Communities, excluding País Vasco and Navarra, which have a *foral* system[13]). Conversely, spending on basic public services per capita in Comunidad Valenciana is below the average of the communities with a common system [Instituto Valenciano de Investigaciones Económicas (IVIE) (2017)].

Two methodological issues regarding regional-level data should be mentioned. First, as gross enrolment ratios are not available for Spanish regions, we use a proxy variable proposed by de la Fuente and Doménech (2015), which is constructed as the proportion of the population for which the maximum level of education achieved is either secondary or tertiary[14]. Second, with respect to government expenditure on education, we derive the value for each region by applying a ratio (the weight of each region's GDP over the total Spanish GDP) to the total expenditure for the country[15]. In addition, the sample ranges from 1971 to 2013 except for education, which goes from 1971 to 2011, and for labour force, which starts in 1977 and goes until 2013. Data are annual in all cases. All variables are used in their logarithmic transformation.

As in the country-level analysis, in a first step we carry out the Ng–Perron unit root tests for the different variables at regional level. As for the aggregate, all variables are used in their first differences[16].

Regarding nonlinearities, we observe how the estimated model including secondary education reflects the dependence of the economic growth on its own past, educational level of the population, physical capital, labour force and (regional) government expenditure on education. Education plays the dual role of being a determining factor of economic growth as well as the force driving its nonlinear behaviour. We also find evidence of asymmetric effects of the variables on economic growth in the regional case; there is a notable difference regarding the sign when comparing the coefficients for the lower and the upper regimes (especially in Castilla y Léon and Cataluña).

Two issues arise in the regional analysis for secondary education. First, this variable is found to be a significant factor for the economic activity at regional level, though it was not statistically significant at country level. Second, overall, the labour force variable does not seem to exert the same influence on economic growth at regional level as it does at country level. The observable relevance of other factors such as physical capital or public expenditure on education might be diminishing the effect of labour force.

Turning to the nonlinear specification for tertiary education, all the variables considered are relevant factors explaining economic growth in the regional analysis. Once more, the educational level plays a key role acting as the transition variable as well as a relevant factor explaining the economic growth. Asymmetries are again found (in particular, for Castilla y León and Cataluña).

One aspect to mention about nonlinear models in the regional analysis is the remarkable influence of physical capital, which is often present in all models, while labour force appears as an explanatory variable to a greater extent in the aggregate analysis than in the disaggregated analysis. Conversely, public spending on education more often appears as an explanatory variable at regional level than at country level. One explanation for this might be that the analysis using the aggregated data could be masking the importance of public spending on education for economic growth in Spain.

Regarding transition functions, one remarkable difference from the aggregate analysis for secondary education is the estimated value for the slope parameter. At regional level, the transition between the lower and the upper regime takes place at a higher speed than in the country-level analysis. The location parameters that determine the thresholds between the extreme regimes are remarkably close to the mean of the transition variable (5-6 per cent), except for Cataluña. Finally, the estimated functions generally display a wide variation range (except for Cataluña).

Regarding tertiary education, the transition between the extreme regimes is smoother at regional than at country level. Nevertheless, the changes between regimes occur at a notable speed. The estimated thresholds between the extreme regimes range from 3 to 6 per cent; these values are very close to the respective enrolment ratio means (mainly around 4 per cent). Thus, the left and the right sides of the logistic function are fairly well balanced in terms of the number of observations. As observed in the secondary education case, there is wide variation in the estimated transition functions. It is worth mentioning the relatively rapid transition between the extreme regimes we observe in the regional analysis. As noted in the country-level analysis, this fact points to the need for threshold specifications, thus reinforcing the relevance of STR models.

Focusing on the evaluation stage, the estimated models for the four regions present no evidence of misspecification following the diagnostic statistics. Furthermore, as in the country analysis, the explanatory power of the nonlinear models substantially outweighs that of the linear specifications.

Finally, just as we did for the aggregate analysis, we examine the residuals of both linear and nonlinear models for the four regions under study. According to the results obtained, the pattern of behaviour at regional level resembles that observed in the country-level analysis, although in some regions (Comunidad Valenciana and Castilla y León) there is even stronger evidence in favour of the nonlinear model.

5. Conclusions and discussion

In this paper, the assumption of linearity in the relationship between education and economic growth is called into question. We opt for a time series framework (analysing the period from 1971 to 2013) and, more specifically, we estimate STR models for Spain. The estimated models reflect how the variation in education (contemporary or one or two periods before) generates nonlinear effects on current economic growth.

We proceed in two stages. First, we focus on the role of secondary and tertiary education in economic growth in Spain (aggregated or country-level analysis). Second, as different solutions emerge in different regions, we perform our nonlinear analysis for a number of representative Spanish regions (disaggregated or regional analysis). We can then test for the existence of regional heterogeneity. Using these two steps, we test whether our initial findings hold up or are sensitive to the level of territorial aggregation used.

According to the obtained results, we find that both the aggregated and disaggregated analyses lead to consistent and coherent results. The linear analysis has shown evidence in favour of a positive correlation between education and economic growth: the higher the educational level (both secondary and tertiary), the higher the economic growth in Spain. When analysing nonlinearities, we observe a number of differences regarding the importance of secondary education as an explanatory variable, as well as the role of other variables. Also, there are some variables that might seem more relevant than they really are as a consequence of statistical aggregation (e.g. labour force, once education is controlled for), while others might seem to be less relevant (e.g. public spending on education).

Both secondary and tertiary education are relevant for economic growth[17]. The obtained results point to the need for a nonlinear model, as there are abrupt shifts between regimes; importantly, we evidence how education plays the dual role of being a determining factor of economic growth as well as the force driving its nonlinear behaviour. However, our aim here is not to provide an answer as to which level of education is more important to boost economic growth in Spain, as we have not focused on comparing the magnitudes of estimated coefficients for different educational levels, and we have not included primary education in our analysis, though this might also be relevant for economic growth in Spain. We leave this important issue for further research.

What we learn from using STR models for the education-economic growth link is that the educational level of the population can be understood as a source of nonlinearities in the economic activity of a country (and of a region). Thus, depending on national and regional educational levels, economic growth behaves differently. By means of these models, we have been able to detect how rapidly economic growth reacts to the variable determining the transition, that is, education.

Notes

1. Our paper is linked to the literature on absorptive capacity as it could be argued that human capital is a key element in a country's absorptive capacity (Cohen and Levinthal, 1990; Zahra and George, 2002, and Márquez-Ramos and Martínez-Zarzoso, 2010). See, for example, the illustration of the role of education in the aerospace sector for establishing sufficient levels of absorptive capacity in Asia (van der Heiden *et al.*, 2015).
2. Spain is an interesting country to explore at regional level because of its high degree of decentralization, through which different solutions emerge in different regions.
3. Under certain circumstances, economic theory or/and data might suggest that a given variable is generated by a nonlinear process. In this sense, it is quite usual for a variable to behave differently according to the values of another variable; for example, it has been shown that the dynamic properties of GDP growth are different in expansion and recession periods. Thus, there would be a change of regimes and the data generating process to be modelled would be a linear process that switches between regimes according to a rule.
4. STR has been mainly used for evidencing asymmetries in economic variables; some leading empirical applications deal with interest rates, unemployment or exchange rates, among other variables. However, the application of STs is not restricted to the economic context. For example, Hall *et al.* (2001) demonstrate how a Smooth Transition Auto Regression (STAR) model fits better than an Autoregressive (AR) specification when describing climate turbulence periods, and Abril-Salcedo *et al.* (2016) estimate the impact of weather conditions on food inflation growth by means of STR models, attempting to account for the fact that the effects of weather anomalies vary over time due to climate change.
5. Note that, for example, in Lucas' model, aggregate output depends on physical capital and human capital (Krueger and Lindahl, 2001).
6. Kalaitzidakis *et al.* (2001) refer to the proportion of government expenditures on education as a measure employed by several researchers for capturing those resources used as inputs into the educational process.
7. Likewise, linearity tests have been carried out and further information on the results is provided in Table AIV (Appendix 3).
8. We are grateful to two anonymous reviewers that helped us to develop this argument.

9. Note that secondary education is found to be significant in the linear model. We anticipate that secondary education is a significant determinant of economic growth in different regions within Spain (see Section 3).
10. Note that each dot in Figure C.1 represents one observation in the sample.
11. The comparatively large value for the standard deviation reflects the numerical difficulties in obtaining an accurate value for γ when it is large, as in this case F functions mimic step functions, and many observations close to the location parameter are needed [see, for example, van Dijk *et al.* (2002)].
12. The specific validation tests do not detect serial dependence in the estimated residuals, there is no need for a second transition function and constancy in the parameters is demonstrated.
13. In a foral system, the autonomous communities are granted the power to maintain, establish and regulate their own tax regimes (as opposed to the common system, where the State is responsible).
14. Note that the results obtained at the national level are not strictly comparable with those obtained in the regional analysis because in the former we use the gross enrolment ratios, whereas in the latter, we use the population for which the maximum level achieved is either secondary or tertiary.
15. Note that we considered a more accurate proxy for regional public spending in education by considering the series provided by the Ministry of Education ("Recursos económicos. Gasto Público", available at www.mecd.gob.es/servicios-al-ciudadano-mecd/estadisticas/educacion/recursos-economicos/gasto-publico.html). However, due to the limitation of the available time series, we rely on the constructed alternative proxy.
16. Full results for the regional analysis are not reported to save space. In Appendix 4, we report the estimated STR models for secondary and tertiary education. The rest of tests and results are available from the authors upon request.
17. The results show that the main difference regarding secondary education in the aggregate and the regional analyses is that this educational level seems to play a more relevant role at a regional scale, where it is both a transition and an explanatory variable.

References

Abril-Salcedo, D.S., Melo-Velandia, L.F. and Parra-Amado, D. (2016), "Impactos de los fenómenos climáticos sobre el precio de los alimentos en Colombia", *Ensayos Sobre Política Económica*, Vol. 34 No. 80, pp. 146-158.

Afzal, M., Rehman, H.U., Farooq, M.S. and Sarwar, K. (2011), "Education and economic growth in Pakistan: a cointegration and causality analysis", *International Journal of Educational Research*, Vol. 50 Nos 5/6, pp. 321-335.

Armellini, M. (2012), "The democratic factor in the education-growth relationship", *Kyklos*, Vol. 65 No. 3, pp. 285-312.

Cheikh, N.B. (2012), "Non-linearities in exchange rate pass-through: evidence from smooth transition models", MPRA Paper no. 39258.

Cohen, D. and Soto, M. (2007), "Growth and human capital: good data, good results", *Journal of Economic Growth*, Vol. 12 No. 1, pp. 51-76.

Cohen, W.M. and Levinthal, D.A. (1990), "Absorptive capacity: a new perspective on learning and innovation", *Administrative Science Quarterly*, Vol. 35 No. 1, pp. 128-152.

Cuestas, J. and Mourelle, E. (2011), "Nonlinearities in real exchange rate determination: do African exchange rates follow a random walk?", *Applied Economics*, Vol. 43 No. 2, pp. 243-258.

De la Fuente, A. (2017), "Series enlazadas de algunos agregados económicos regionales, 1955-2014. Parte I: metodología", VAB, PIB y puestos de trabajo. RegDat Versión 5.0-I. FEDEA, Estudios sobre la Economía Española no. 2017-08.

De la Fuente, A. and Doménech, R. (2015), "El nivel educativo de la población en españa y sus regiones (1960-2011)", [RegDatEdu_v40_1960_2011]. FEDEA and BBVA Research, Estudios sobre la Economía Española no. 2015/05.

Duflo, E. (2001), "Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment", *American Economic Review*, Vol. 91 No. 4, pp. 795-813.

Eitrheim, Ø. and Teräsvirta, T. (1996), "Testing the adequacy of smooth transition autoregressive models", *Journal of Econometrics*, Vol. 74 No. 1, pp. 59-75.

Granger, C.W.J. and Teräsvirta, T. (1993), *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.

Hall, A.D., Skalin, J. and Teräsvirta, T. (2001), "A nonlinear time series model of El niño", *Environmental Modelling and Software*, Vol. 16 No. 2, pp. 139-146.

Hanushek, E.A. and Wößmann, L. (2010), "Education and economic growth", in Peterson, P., Baker, E. and McGaw, B. (Eds), *International Encyclopedia of Education*, Vol. 2, Elsevier, Oxford, pp. 245-252.

Instituto Valenciano de Investigaciones Económicas (IVIE) (2017), "Insuficiencia de recursos Para servicios públicos fundamentales", Papeles de Financiación Autonómica no. 2, November 2017.

Jalil, A. and Idrees, M. (2013), "Modeling the impact of education on the economic growth: evidence from aggregated and disaggregated time series data of Pakistan", *Economic Modelling*, Vol. 31, pp. 383-388.

Kalaitzidakis, P., Mamuneas, T.P., Savvides, A. and Stengos, T. (2001), "Measures of human capital and nonlinearities in economic growth", *Journal of Economic Growth*, Vol. 6 No. 3, pp. 229-254.

Krueger, A.B. and Lindahl, M. (2001), "Education for growth: why and for whom?", *Journal of Economic Literature*, Vol. 39 No. 4, pp. 1101-1136.

Márquez-Ramos, L. and Martínez-Zarzoso, I. (2010), "The effect of technological innovation on international trade", *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 4 Nos 2010/11, pp. 1-37.

Mejía-Reyes, P., Osborn, D.R. and Sensier, M. (2010), "Modelling real exchange rate effects on output performance in Latin America", *Applied Economics*, Vol. 42 No. 19, pp. 2491-2503.

Ng, S. and Perron, P. (2001), "Lag length selection and the construction of unit root tests with good size and power", *Econometrica*, Vol. 69 No. 6, pp. 1519-1554.

Öcal, N. and Osborn, D.R. (2000), "Business cycle non-linearities in UK consumption and production", *Journal of Applied Econometrics*, Vol. 15 No. 1, pp. 27-43.

Potter, S.M. (1999), "Nonlinear time series modelling: an introduction", *Journal of Economic Surveys*, Vol. 13 No. 5, pp. 505-528.

Self, S. and Grabowski, R. (2004), "Does education at all levels cause growth? India, a case study", *Economics of Education Review*, Vol. 23 No. 1, pp. 47-55.

Sensier, M., Öcal, N. and Osborn, D.R. (2002), "Asymmetric interest rate effects for the UK real economy", *Oxford Bulletin of Economics and Statistics*, Vol. 64 No. 4, pp. 315-339.

Sunde, U. and Vischer, T. (2015), "Human Capital and growth: specification matters", *Economica*, Vol. 82 No. 326, pp. 368-390.

Teräsvirta, T. (1994), "Specification, estimation, and evaluation of smooth transition autoregressive models", *Journal of the American Statistical Association*, Vol. 89 No. 425, pp. 208-218.

Teräsvirta, T. (1998), "Modeling economic relationships with smooth transition regressions", in Ullah, A. and Giles, D.E.A. (Eds), *Handbook of Applied Economic Statistics*, Marcel Dekker, New York, NY, pp. 507-552.

Teräsvirta, T. and Anderson, H.M. (1992), "Characterizing nonlinearities in business cycles using smooth transition autoregressive models", *Journal of Applied Econometrics*, Vol. 7, pp. S119-136.

van der Heiden, P., Pohl, C., Mansor, S.B. and van Genderen, J. (2015), "The role of education and training in absorptive capacity of international technology transfer in the aerospace sector", *Progress in Aerospace Sciences*, Vol. 76, pp. 42-54.

van Dijk, D., Franses, P.H. and Lucas, A. (1999), "Testing for smooth transition nonlinearity in the presence of outliers", *Journal of Business and Economic Statistics*, Vol. 17 No. 2, pp. 217-235.

van Dijk, D., Teräsvirta, T. and Franses, P.H. (2002), "Smooth transition autoregressive models-a survey of recent developments", *Econometric Reviews*, Vol. 21 No. 1, pp. 1-47.

Zahra, S.A. and George, G. (2002), "Absorptive capacity: a review, reconceptualization, and extension", *Academy of Management Review*, Vol. 27 No. 2, pp. 185-203.

Corresponding author
Estefanía Mourelle can be contacted at: estefania.mourelle@udc.es

Appendix 1

Smooth Transition (ST) models belong to the family of state-dependent models where the data-generating process is linear but switches between a certain number of regimes according to a rule (see, for example, Teräsvirta, 1994, 1998, and Potter, 1999). This parameterisation allows for capturing different types of behaviour that a linear model cannot appropriately characterise; once the state is given, the model is locally linear, involving an easy interpretation of the local dynamics. In contrast to other regime-switching models (such as Markov–Switching or threshold models), STs consider that the change between regimes is smooth over time, rather than abrupt, which is normally a more realistic situation; in any case, STs nest some threshold models as particular cases.

In recent decades, STs have been a popular choice for economic time series, performing well in capturing cyclical behaviour in macroeconomic variables (Teräsvirta and Anderson, 1992; Mejía-Reyes *et al.*, 2010; Cuestas and Mourelle, 2011, among others). For further details on these models, see Teräsvirta (1994, 1998) and van Dijk *et al.* (2002).

In its basic version, the regime-switching STR specification considers two distinct regimes corresponding to F = 0 and F = 1, and the transition from one regime to the other is smooth over time, meaning that parameters in (1) gradually change with the transition variable. Two formulations are generally used for F: the logistic and the exponential function. In the logistic model, the extreme regimes are associated with s_t values far above or below c, where the dynamics may be different; conversely, in the exponential case, the extreme regimes are related to low and high absolute values of s_t, with rather similar dynamics, which may differ in the transition period.

With regards the modelling procedure, the STR modelling cycle has traditionally relied on the iterative methodology proposed by Teräsvirta (1994) for the univariate case: specification, estimation and evaluation of the model. The usual starting point is identifying the linear model that best characterises the behaviour of the series under study; once this specification has been determined, its suitability regarding the relationship being analysed is tested. If the null hypothesis of a linear process against the alternative of an STR one is rejected, a preliminary specification of the nonlinear model is defined. Then, the parameters of the STR specification are estimated by nonlinear least squares.

Nevertheless, an important branch of the literature on this specification does not follow the above procedure so strictly, as, among other issues, linearity tests suffer from size and power problems under certain circumstances (van Dijk *et al.*, 2002). In this sense, it is argued that it is possible to develop nonlinear formulations that improve the fit of the linear ones without having to do the previous tests, i.e. the data themselves would reveal the potential existence of nonlinearities; the validation process will determine whether or not the model has correctly captured the nonlinear behaviour (Granger and Teräsvirta, 1993).

This alternative procedure is done through an extensive search of STR models by defining a grid for (γ, c). This is the strategy proposed by several authors, such as Öcal and Osborn (2000), or Sensier *et al.* (2002), among others, and the one we adopt in our study; we try different values of γ and use the sample mean of the transition variable for c. Where parameter convergence is reached, models are subjected to further refinement; cross-parameter restrictions are evaluated to increase efficiency and non-significant coefficients are dropped to conserve degrees of freedom.

As mentioned above, less emphasis is given to the initial stages of the modelling process and more attention is paid to the validation of the estimated model, as it will reveal any possible specification inadequacy (van Dijk *et al.*, 2002). Most of the tests commonly applied to dynamic models are also valid in the STR framework. Besides, STR estimation is based on the assumption of no residual autocorrelation and parameter constancy, making it necessary to test these hypotheses. Eitrheim and Teräsvirta (1996) develop several evaluation tests especially derived for ST models, such as the test of residual serial independence against processes of different orders and the test of parameter constancy against changing parameters under the alternative, which refer to the two previous assumptions. In addition, these authors define the test of no remaining nonlinearity in the residuals.

Appendix 2

Variable name	Definition	Source: Spain	Source: Regions
Gross Domestic Product (GDP)	Real Gross Domestic Product (as index number)	International Monetary Fund's International Financial Statistics (IFS) database	FEDEA, reference: de la Fuente (2017)
Education, secondary education (ENRS)	Gross Enrolment Ratios for secondary education	World Development Indicators, World Bank	
	Proportion of population with secondary education		FEDEA, reference: de la Fuente and Doménech (2015)
Education, tertiary education (ENRT)	Gross Enrolment Ratios for tertiary education	World Development Indicators, World Bank	
	Proportion of population with tertiary education		FEDEA, reference: de la Fuente and Doménech (2015)
Labour force (LABS, LABT)	Proportion of active population with secondary (tertiary) education over total active population	The Valencian Institute of Economic Research (IVIE, in its Spanish acronym)	The Valencian Institute of Economic Research (IVIE, in its Spanish acronym)
Physical capital (PHY)	Real Gross Fixed Capital Formation (as index number)	The Valencian Institute of Economic Research (IVIE, in its Spanish acronym)	The Valencian Institute of Economic Research (IVIE, in its Spanish acronym)
Government expenditure (EXP)	Government expenditure on education (as a percentage of GDP)	World Development Indicators, World Bank	World Development Indicators, World Bank
	Government expenditure on education in real (2010) terms		Own calculation using World Development Indicators (World Bank) and FEDEA

Source: Own elaboration

Table AI.
Data set description

Low government expenditure per capita	High government expenditure per capita
Low income per capita	
Andalucía	Canarias
Galicia	Extremadura
Castilla-la-Mancha	Cantabria
Castilla y León	Comunidad Valenciana
Asturias	Murcia
High income per capita	
Aragón	Navarra
La Rioja	País Vasco
Madrid	
Cataluña	
Baleares	

Note: Regions are classified into four groups as follows: regions are ordered from highest to lowest income levels (GDP per capita, average 2000-2013) and from highest to lowest government expenditure on education per capita (average 2000-2013), then a threshold is constructed for both GDP and government expenditure by calculating the average of the sample in each magnitude
Source: Own elaboration

Table AII. Classification matrix

Appendix 3

Variable	MZα	MZt	MSB	MPt
$lGDP_t$	−12.764	−2.482	0.194	7.382
$lENRS_t$	−4.395	−1.464	0.333	20.569
$lENRT_t$	−3.210	−1.065	0.332	24.209
$lPHY_t$	−18.808	−2.948	0.157	5.550
$lLABS_t$	−12.312	−2.283	0.185	8.436
$lLABT_t$	−5.759	−1.696	0.295	15.822
$lEXP_t$	−2.645	−0.910	0.344	26.870
Significance level				
1%	−23.8000	−3.42000	0.14300	4.03000
5%	−17.3000	−2.91000	0.16800	5.48000
10%	−14.2000	−2.62000	0.18500	6.67000

Note: Logarithm of original variables used for calculations (l). The asymptotic critical values are:
Source: Own elaboration

Table AIII. Ng–Perron unit root test results

Table AIV. Linearity tests against logistic smooth transition regression (p-values). country-level analysis

Transition variable	Secondary education	Tertiary education
$ENRS/T_t$	0.0038	0.0022
$ENRS/T_{t-1}$	0.5345	0.0382
$ENRS/T_{t-2}$	0.7189	0.1960

Notes: All variables are first differences of the logarithmic transformation, The transition variable is assumed to be the difference of (the logarithm of) the gross enrolment ratio up to two lags, We use tests robust to outliers because our sample could be affected by the "baby boom" in Spain in the 60 s, which might potentially be reflected in the education variables (i.e. our regressor of interest). Standard linearity tests (Teräsvirta, 1994, 1998) can be misleading in the presence of outliers. van Dijk *et al.* (1999) advocate linearity tests that display a better level and power performance than the standard ones if outliers are present
Source: Own elaboration

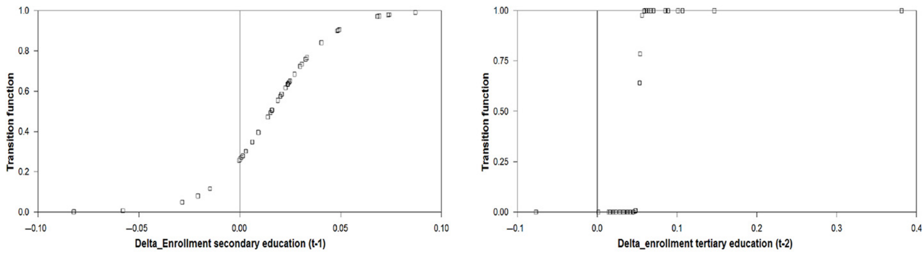
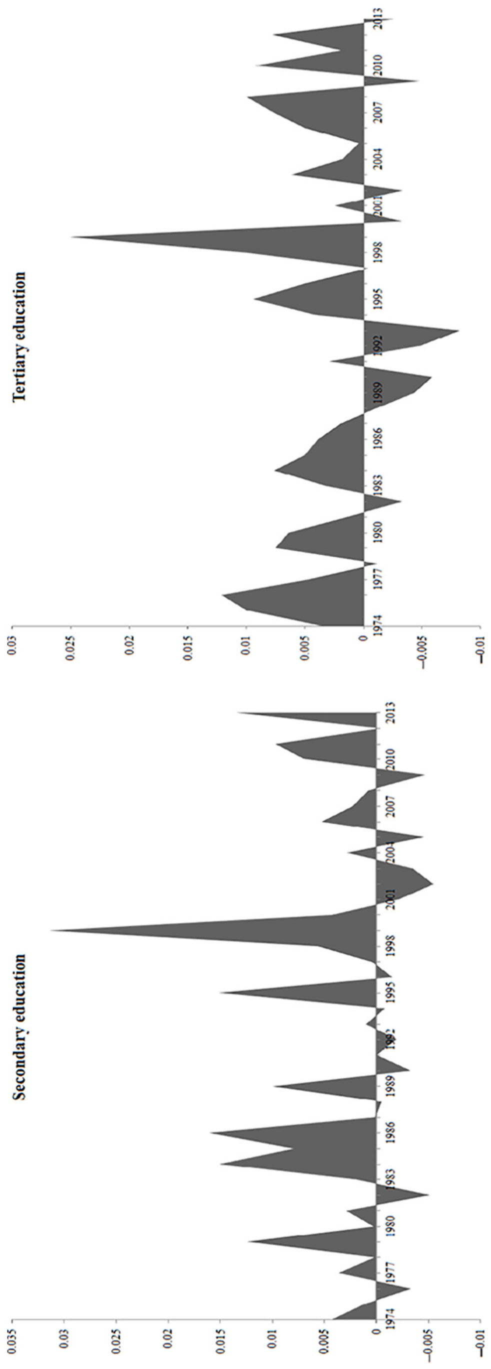


Figure A1. Estimated transition functions

Source: Own elaboration (Country-level analysis)



Source: Own elaboration (Country-level analysis)

Figure A2.
Differences between the absolute values of the estimated residuals (linear and STR model)

Appendix 4

Table AV. Estimated STR models for secondary education (Regional analysis)

	CASTILLA Y LEÓN		CATALUÑA		COM. VALENCIANA		PAÍS VASCO	
	Linear model	STR model	Linear model	STR model	Linear model	STR model	Linear model	STR model
		Transition variable: $ENRS_t$		Transition variable: $ENRS_t$		Transition variable: $ENRS_{t-1}$		Transition variable: $ENRS_t$
GDP_{t-1}	0.43 (0.10)			1.07 (0.28)	0.38 (0.09)	−0.46 (0.10)	0.50 (0.14)	−0.63 (0.16)
GDP_{t-2}								
$ENRS_t$		0.40 (0.15)		5.86 (1.72)				1.24 (0.36)
$ENRS_{t-1}$				−0.51 (0.28)		0.74 (0.16)		
$ENRS_{t-2}$			0.18 (0.06)			0.16 (0.05)	0.08 (0.04)	
PHY_t		0.10 (0.02)	0.08 (0.03)		0.12 (0.03)	0.10 (0.02)		
PHY_{t-1}		−0.09 (0.01)	0.07 (0.03)	0.39 (0.10)				
PHY_{t-2}							0.10 (0.02)	
$LABS_t$								
$LABS_{t-1}$						0.11 (0.03)		
$LABS_{t-2}$				−0.36 (0.17)				
EXP_t	0.16 (0.03)	0.46 (0.06)	0.17 (0.03)		0.17 (0.03)	0.59 (0.06)	0.19 (0.04)	0.80 (0.08)
EXP_{t-1}				−1.23 (0.41)			−0.11 (0.05)	
EXP_{t-2}		−0.25 (0.04)				−0.43 (0.07)		−0.31 (0.07)
$F(s_t)\times GDP_{t-1}$				−1.20 (0.40)		0.47 (0.18)		0.63 (0.16)
$F(s_t)\times GDP_{t-2}$		2.14 (0.35)				−0.77 (0.22)		−1.57 (0.38)
$F(s_t)\times ENRS_t$		−0.44 (0.16)		−5.66 (1.70)		−1.73 (0.26)		−1.80 (0.54)
$F(s_t)\times ENRS_{t-1}$								−1.27 (0.42)
$F(s_t)\times ENRS_{t-2}$						0.16 (0.05)		
$F(s_t)\times PHY_t$		−0.39 (0.07)		0.06 (0.02)		−0.21 (0.04)		
$F(s_t)\times PHY_{t-1}$		−0.09 (0.01)		−0.37 (0.10)				0.15 (0.04)
$F(s_t)\times PHY_{t-2}$								
$F(s_t)\times LABS_t$		0.72 (0.14)						
$F(s_t)\times LABS_{t-1}$		−0.57 (0.13)				0.11 (0.03)		
$F(s_t)\times LABS_{t-2}$				0.48 (0.20)		0.25 (0.06)		0.95 (0.22)
$F(s_t)\times EXP_t$		−0.46 (0.06)		0.22 (0.07)				
$F(s_t)\times EXP_{t-1}$		0.20 (0.07)		1.42 (0.44)				
$F(s_t)\times EXP_{t-2}$		−0.25 (0.04)				0.51 (0.14)		0.63 (0.18)
Γ		5.37 (2.01)		4.60 (2.21)		8.99 (3.64)		12.88 (5.02)
C		0.06 (0.00)		0.03 (0.01)		0.06 (0.00)		0.05 (0.00)
$\overline{R}^2$		0.89	0.63	0.87	0.69	0.94	0.46	0.85

(*continued*)

Table AV.

	CASTILLA Y LEÓN		CATALUÑA		COM. VALENCIANA		PAÍS VASCO	
	Linear model	STR model	Linear model	STR model	Linear model	STR model	Linear model	STR model
		Transition variable: $ENRS_t$		Transition variable: $ENRS_t$		Transition variable: $ENRS_{t-1}$		Transition variable: $ENRS_t$
s^2/s^2_L		0.12		0.16		0.08		0.13
AUTO		3.00 (0.10)		5.87 (0.03)		3.24 (0.10)		2.47 (0.14)
NL		0.23 (0.93)		0.28 (0.96)		0.33 (0.89)		1.85 (0.40)
PC		0.87 (0.64)		1.78 (0.54)		0.62 (0.74)		2.87 (0.21)

Notes: All variables are first differences of the logarithmic transformation. Values after regression coefficients are SEs of the estimates; $-R^2$ is the adjusted coefficient of determination; s^2/s^2_L is the variance ratio of the residuals from the STR model and the linear regression; AUTO is the test for residual autocorrelation of order 1; NL is the test for no remaining nonlinearity; PC is the general parameter constancy test. Numbers in parentheses after misspecification tests statistics are p-values. This table reports the estimated models for secondary and tertiary education using the explanatory variables listed in the first column. For each educational level, the first column presents the linear model and the second column is for the STR specification. Regarding the nonlinear model, the upper part of the table corresponds to the "lower" regime (F = 0), and the next part introduces the interactions with the transition function [equation (1)]. The smoothness of the transition between regimes is given by the slope parameter γ and c represents the estimated threshold

Source: Own elaboration

Table AVI. Estimated STR models for tertiary education (Regional analysis)

	CASTILLA Y LEÓN		CATALUÑA		COM. VALENCIANA		PAÍS VASCO	
	Linear model	STR model Transition variable: $ENRT_{t-2}$	Linear model	STR model Transition variable: $ENRT_{t-1}$	Linear model	STR model Transition variable: $ENRT_t$	Linear model	STR model Transition variable: $ENRT_{t-1}$
GDP_{t-1}	0.47 (0.14)						0.29 (0.10)	−0.47 (0.13)
GDP_{t-2}	−0.26 (0.12)	−1.21 (0.20)		0.24 (0.11)				−0.36 (0.11)
$ENRT_t$		0.37 (0.13)		−0.47 (0.27)	0.44 (0.23)	1.14 (0.13)	−0.83 (0.27)	−0.71 (0.26)
$ENRT_{t-1}$		−0.91 (0.19)			−1.08 (0.30)	−0.70 (0.07)		0.75 (0.34)
$ENRT_{t-2}$	0.14 (0.08)			0.69 (0.37)	0.79 (0.22)		1.17 (0.28)	
PHY_t	0.06 (0.03)	0.04 (0.02)	0.08 (0.02)	0.07 (0.03)	0.12 (0.02)	0.15 (0.01)	0.10 (0.02)	0.07 (0.02)
PHY_{t-1}		−0.11 (0.02)		0.05 (0.02)		0.07 (0.01)		0.09 (0.03)
PHY_{t-2}		0.06 (0.02)			0.08 (0.02)	0.04 (0.02)	0.05 (0.02)	0.09 (0.02)
$LABT_t$				−0.15 (0.08)		0.09 (0.01)		
$LABT_{t-1}$		−0.16 (0.03)	0.10 (0.04)			0.06 (0.02)		0.10 (0.02)
$LABT_{t-2}$					0.11 (0.05)	0.06 (0.01)		
EXP_t	0.21 (0.03)	0.33 (0.04)	0.27 (0.03)	0.24 (0.04)	0.29 (0.07)	0.11 (0.04)	0.20 (0.04)	0.25 (0.07)
EXP_{t-1}	−0.07 (0.04)	0.53 (0.09)			−0.13 (0.06)	−0.25 (0.04)	−0.22 (0.04)	
EXP_{t-2}						0.19 (0.03)		
$F(s_t)\times GDP_{t-1}$		0.98 (0.21)				−0.70 (0.14)		
$F(s_t)\times GDP_{t-2}$		0.69 (0.22)		0.24 (0.11)				
$F(s_t)\times ENRT_t$		−1.65 (0.72)		0.96 (0.31)		−0.50 (0.12)		
$F(s_t)\times ENRT_{t-1}$		2.38 (0.72)				−0.70 (0.07)		−2.49 (0.53)
$F(s_t)\times ENRT_{t-2}$				−0.69 (0.37)				
$F(s_t)\times PHY_t$		−0.09 (0.04)		−0.07 (0.03)				0.17 (0.06)
$F(s_t)\times PHY_{t-1}$		0.25 (0.04)		0.05 (0.02)				
$F(s_t)\times PHY_{t-2}$		−0.19 (0.05)		−0.26 (0.06)		0.17 (0.02)		
$F(s_t)\times LABT_t$				0.21 (0.10)		0.09 (0.01)		
$F(s_t)\times LABT_{t-1}$				0.21 (0.05)				0.10 (0.02)
$F(s_t)\times LABT_{t-2}$						0.06 (0.01)		0.30 (0.05)
$F(s_t)\times EXP_t$		0.33 (0.04)				0.59 (0.10)		0.93 (0.22)
$F(s_t)\times EXP_{t-1}$		−0.96 (0.10)				0.58 (0.08)		0.36 (0.16)
$F(s_t)\times EXP_{t-2}$				−0.46 (0.12)		−0.37 (0.08)		0.63 (0.19)
Γ		12.67 (5.67)		22.16 (20.26)		5.71 (1.31)		5.38 (1.29)
C		0.04 (0.00)		0.04 (0.00)		0.04 (0.00)		0.04 (0.00)
$-R^2$	0.70	0.94	0.76	0.82	0.85	0.98	0.75	0.90
s^2/s^2_L		0.08		0.39		0.06		0.17
AUTO		4.63 (0.06)		2.67 (0.12)		2.13 (0.18)		0.85 (0.38)

(*continued*)

	CASTILLA Y LEÓN		CATALUÑA		COM. VALENCIANA		PAÍS VASCO	
	Linear model	STR model Transition variable: $ENRT_{t-2}$	Linear model	STR model Transition variable: $ENRT_{t-1}$	Linear model	STR model Transition variable: $ENRT_t$	Linear model	STR model Transition variable: $ENRT_{t-1}$
NL		1.42 (0.48)		0.37 (0.88)		0.58 (0.78)		0.46 (0.83)
PC		3.00 (0.28)		3.31 (0.26)		4.80 (0.34)		0.50 (0.82)

Notes: All variables are first differences of the logarithmic transformation. Values after regression coefficients are SEs of the estimates; $\overline{R}^2$ is the adjusted coefficient of determination; s^2/s^2_L is the variance ratio of the residuals from the STR model and the linear regression; AUTO is the test for residual autocorrelation of order 1; NL is the test for no remaining nonlinearity; PC is the general parameter constancy test. Numbers in parentheses after misspecification tests statistics are *p*-values. This table reports the estimated models for secondary and tertiary education using the explanatory variables listed in the first column. For each educational level, the first column presents the linear model and the second column is for the STR specification. Regarding the nonlinear model, the upper part of the table corresponds to the "lower" regime ($F = 0$), and the next part introduces the interactions with the transition function [equation (1)]. The smoothness of the transition between regimes is given by the slope parameter γ and c represents the estimated threshold
Source: Own elaboration

Table AVI.